

## Test de Estadística Descriptiva

### Probabilidad y Estadística I. Marzo 2013.

Realizar el test en esta hoja. **Duración:** 55 minutos. Tanto el alumno que copie como el que se deje copiar, no podrá examinarse hasta el próximo curso.

---

Apellidos:

Nombre:

---

1. (1 punto) En una Escuela de natación se recogen, entre otras, las siguientes características de cada alumno. Clasifica cada característica como nominal, ordinal, cuantitativa discreta o cuantitativa continua.
  - a) Número de días que falta a entrenar. **Cuantitativa discreta.**
  - b) Profesor de natación de su grupo. **Nominal.**
  - c) Tiempo que tarda en nadar los 100 m. espalda. **Cuantitativa continua.**
  - d) Lugar que ocupa según su tiempo en 100 m. libres. **Ordinal.**
  - e) Número de alumnos en su grupo. **Cuantitativa discreta.**
2. (1 punto) Indica si son verdaderas o falsas las siguientes afirmaciones:
  - a) La media aritmética es una medida de posición central y en su cálculo interviene el valor de todos los datos. **Verdadero.**
  - b) El coeficiente de correlación lineal de Pearson entre dos características X e Y y las varianzas de X e Y siempre tienen el mismo signo. **Falso.**
  - c) La mediana y la media aritmética nunca coinciden, pero aportan información complementaria de la distribución. **Falso.**
  - d) La covarianza entre dos características, al igual que su coeficiente de correlación lineal, toma siempre valores entre -1 y 1. **Falso.**
  - e) El valor de la media aritmética siempre está comprendido entre el menor y el mayor valor observado. **Verdadero.**
  - f) La media aritmética es menos sensible a los valores extremos o atípicos que la mediana. **Falso.**
  - g) El coeficiente de correlación lineal de Pearson y el coeficiente de variación siempre tienen el mismo signo. **Falso.**
  - h) La desviación típica de una característica se expresa en las mismas unidades de medida que la característica. **Verdadero.**
  - i) Si en un conjunto de datos se verifica que  $\sum_{i=1}^n (x_i - \bar{x})^3 < 0$ , entonces el coeficiente de asimetría es positivo. **Falso.**
  - j) La mediana es menos sensible a los valores extremos o atípicos que la media. **Verdadero.**



3. (1 punto) Completa los huecos en las siguientes frases:

- Si no hay datos atípicos, el bigote derecho de un diagrama de caja llegará hasta **el máximo**.
- Si la mediana de una serie de datos es 64, la frecuencia relativa acumulada del decil 7 ( $D_7$ ) es **0.7**.
- El cuartil **2**, el decil **5**, el percentil **50** y la mediana son el mismo valor.
- Si quiero comparar la dispersión de dos conjuntos de datos medidos en unidades diferentes, debo usar **el coeficiente de variación**. En cambio, no debería usar **la varianza (o desviación típica)** ya que depende de las unidades de medida.
- Cuando existen datos atípicos, no es adecuado el empleo de la **media** como medida de tendencia central.

4. (2 puntos) Dada la siguiente tabla de frecuencias conjunta de las variables  $X = \text{duración de las consultas médicas (en minutos)}$  e  $Y = \text{especialidad médica}$  de las últimas 79 consultas en un Centro de Salud:

X   Y	Medicina General	Pediatría	Cardiología	Psiquiatría	Marg. X
[0, 15)	10	<b>2</b>	0	0	<b>12</b>
[15,30)	12	15	<b>5</b>	1	<b>33</b>
[30,45)	<b>1</b>	3	6	<b>12</b>	<b>22</b>
[45,60]	1	0	1	10	<b>12</b>
Marg. Y	24	20	<b>12</b>	23	<b>79</b>

- Completar la tabla.
- ¿Qué tipo de gráfico sería adecuado para representar la variable  $Y = \text{especialidad médica}$ ? **Diagrama de barras o de sectores.**
- Dar la tabla de frecuencias de la distribución de la duración de las consultas médicas dentro de la especialidad de Cardiología. ¿Cuál es la duración media de las consultas de Cardiología?
- ¿Cuál es la moda de la variable  $Y = \text{especialidad médica}$ ? **Medicina General.**

X Y=Cardio	$x'_i$	$n_i$	$f_i$	$N_i$	$F_i$
[15,30)	22.5	5	0.42	5	0.42
[30,45)	37.5	6	0.5	11	0.92
[45,60]	52.5	1	0.08	12	1
TOTAL		12	1		

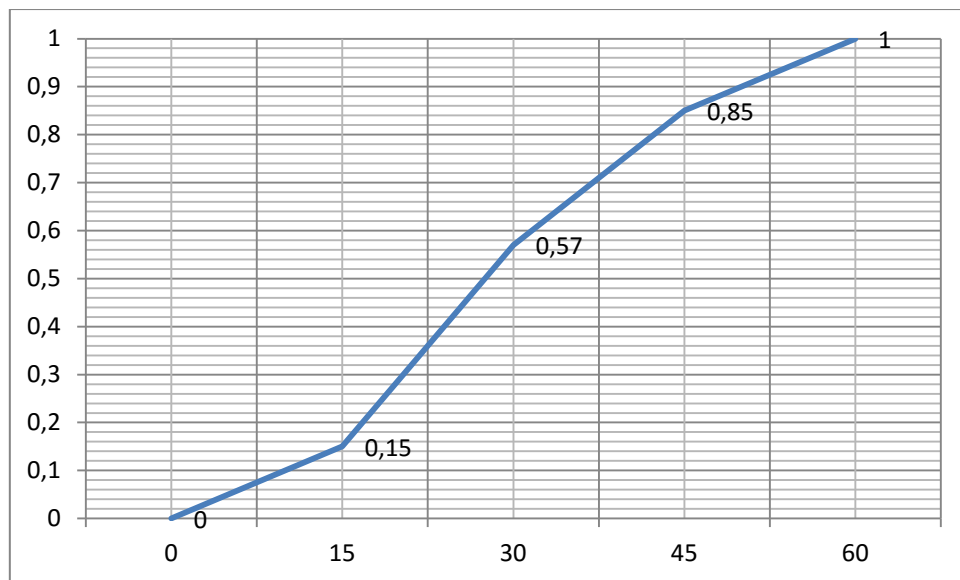
$$\overline{x|(Y = \text{cardio})} = \frac{(22.5) \cdot 5 + (37.5) \cdot 6 + (52.5) \cdot 1}{12} = \frac{390}{12} = 32.5$$



5. (1 punto) Construir la tabla (completa) de la distribución de frecuencias de la variable  $X = \text{duración de las consultas del ejercicio anterior}$ .

$X Y=\text{Cardio}$	$x'_i$	$n_i$	$f_i$	$N_i$	$F_i$
[0, 15)	7.5	12	0.15	12	0.15
[15,30)	22.5	33	0.42	45	0.57
[30,45)	37.5	22	0.28	67	0.85
[45,60]	52.5	12	0.15	79	1
<b>TOTAL</b>		79	1		

6. (2 puntos) Dibujar en el gráfico el polígono de frecuencias acumuladas de la distribución de  $X$  del ejercicio anterior de forma que se aprecien claramente los puntos representados.



- a) Calcula la duración máxima del 35% de las consultas que menos duraron.  
b) Calcula el percentil 80 de la distribución de frecuencias de la variable  $X = \text{duración de las consultas}$ .

a) Piden el  $P_{35}$ , que está en el intervalo [15,30).

$$P_{35} = 15 + \frac{(0.35) \cdot 79 - 12}{33} \cdot 15 = 22.113$$

Interpolando sale  $P_{35} = 22.14$

b) Piden el  $P_{80}$  que está en el intervalo [30,45).

$$P_{80} = 30 + \frac{(0.8) \cdot 79 - 45}{22} \cdot 15 = 42.409$$



Interpolando sale  $P_{80} = 42.32$

7. (1 punto) La distribución de frecuencias de una variable  $X$  tiene media  $\bar{x} = 275$  y varianza  $m_2 = 2500$ . Se quiere una variable, definida a partir de  $X$ , que tenga media 5 y varianza 1. ¿Qué transformación lineal debe aplicarse a los valores de  $X$ ?

$$\text{Si } Y = aX + b, \quad \bar{Y} = a\bar{X} + b, \quad m_{2Y} = a^2 m_{2X}$$

Debe ser:

$$\bar{Y} = a\bar{X} + b = 275a + b = 5$$

$$a^2 m_{2X} = a^2 2500 = 1$$

De donde:

$$a = \frac{1}{50}, \quad b = -\frac{1}{2}$$

Y la transformación será:

$$Y = \frac{1}{50} X - \frac{1}{2}$$

8. (1 punto) Una empresa se plantea conceder una ayuda anual a sus trabajadores con hijos que consistirá en 300 euros anuales por hijo. Si este colectivo tiene una media de 2 hijos con varianza igual a 0.25, el coeficiente de variación de la variable  $Y = \text{ayuda económica por trabajador}$  será:

- a) 0.25. Correcta.
- b) 0.0008
- c) 75
- d) 0.2